Genetic testing in beef practice

Jason B. Osterstock, DVM, PhD
Zoetis, 333 Portage St., Kalamazoo, MI 49007

Abstract

Genomic technologies are emerging as an important tool in beef genetic improvement strategies. Genetic tests are used extensively by many progressive seedstock breeders to increase accuracy of selection and accelerate genetic progress. Increasing use of these technologies in the seedstock sector have also led to development of applications for commercial cattle. There are several keys to technical evaluation of genomic technologies. The first is the accuracy of the prediction which can be expressed as a genetic correlation or an increase in accuracy of EPDs when incorporated into genetic evaluation. The second is the association between the prediction and the observed phenotype. In the case of the genomically enhanced EPD, this is implicit in the reported accuracy. For commercial applications, this should include some external validation of the predictions in a population independent of that used for the development of the test. Provided these validation criteria are met, the application of the technology is very straightforward and effectively the same as historical practices. The principal differences are that the selection decisions can be made earlier in life with greater accuracy, including for traits that are not expressed phenotypically until much later in life. This accelerates genetic progress by minimizing selection mistakes and allowing producers to identify and exploit superior genetics much sooner than with traditional approaches.

Key words: cattle, beef, genomics, EPD

Résumé

Les technologies de la génomique deviennent des outils de plus en plus importants dans le contexte des stratégies d'amélioration de la génétique des bovins de boucherie. Les tests génétiques sont utilisés par plusieurs producteurs de géniteurs avant-gardistes afin d'augmenter la précision de la sélection et d'accélérer le progrès génétique. L'utilisation accrue de ces technologies dans la filière des géniteurs a aussi permis le développement d'applications pour les bovins mis en marché. Il y a plusieurs éléments clés dans l'évaluation technique des technologies de la génomique. Le premier est la précision de la prédiction qui peut s'exprimer par la corrélation génétique ou par l'augmentation du degré de précision des écarts prévus de la descendance (EPD) lorsqu'incorporés dans l'évaluation génétique. Le second est l'association entre le phénotype prédit et observé. Dans le cas des EPD rehaussés génomiquement, ceci découle de la précision rapportée. Pour les applications commerciales, il est important d'avoir une validation externe des prédictions dans une population indépendante de celle utilisée pour le développement du test. En autant que ces critères de validation soient rencontrés, l'application de cette technologie est directe et ressemble en fait aux pratiques antérieures. La principale différence réside dans le fait que les décisions de sélection peuvent se faire plus tôt dans la vie et avec plus de précision incluant les caractéristiques qui ne s'expriment phénotypiquement que plus tard la vie. Ceci accélère le progrès génétique en minimisant les erreurs de sélection et en permettant aux producteurs d'identifier et d'exploiter des génétiques de plus haut niveau plus tôt qu'avec des approches traditionnelles.

Introduction

The use of genetic testing in beef production has made significant advancements in the last 5 years. This has included expanded scope of traits that can be predicted from genomic data, increased accuracy of the resulting predictions, and greater appreciation of the benefits associated with use of the technology. Adoption of genetic testing has increased accordingly with the greatest growth observed in the seedstock sector as a complement to existing genetic evaluation systems, although an increasing number of applications for commercial cattle are also becoming available.

The significance of these developments to beef practitioners will vary considerably depending upon their role. In some instances, they may be very involved, providing assistance in analysis and interpretation of results for their customers, perhaps very consultative in their contribution. In other instances, they may serve principally as a trusted advisor that can help producers evaluate the decision of whether to apply the technology in their herds. Unfortunately, the technology moves very quickly, requiring a good foundation of basic knowledge about genomic technologies that can readily be adapted as new applications emerge.

Drivers of Genetic Progress

Genetic progress, whether in seedstock or commercial production, is influenced by the same principles. The simplest approach is to examine the classical genetic progress formula:

$$\Delta G = \frac{(r \cdot i \cdot \sigma)}{GI}$$

Where ΔG is the increase in average genetic merit; r is the accuracy of the prediction of genetic merit; i is the selection intensity; σ is the genetic variation of the trait under selec-

tion; and *GI* is the generation interval defined as the average age of parents.

Genomic technologies can influence ΔG in 2 primary ways. The first and most important is the increase in accuracy of the genetic predictions. As accuracy increases, genetic progress increases. The second is by reducing generation interval. In some instances, the availability of higher accuracy data from genomic technologies may allow superior animals to be used as breeding animals more aggressively earlier in life, as may be the case for bulls used for artificial insemination or identifying donor females at an earlier age. Decreasing the average age of parents also increases the rate of genetic progress.

Defining Accuracy

Without a doubt, the most challenging aspect of becoming comfortable with evaluating genomic technologies is understanding how to assess accuracy of the predictions of genetic merit. Some of this is related to the terminology applied. Some is just wrapping one's head around the statistical elements involved in estimating accuracy. Regardless of the root of the confusion, there are some common metrics used to describe accuracy, and some practical context which can be used to better understand what it means.

The accuracy of a prediction of genetic merit is an assessment of how well that prediction reflects an animal's true, but unknown, genetic merit for a given trait. The most common statistic used in quantitative genetics to describe accuracy is the genetic correlation. This is a parameter estimated in genetic analyses that describes the correlation between a predictor and an estimate of genetic potential, also referred to as a breeding value. This correlation varies between 0 and 1 with 1 representing a perfect prediction. Generally, higher accuracy is achieved as more information contributes to the prediction, and for more highly heritable traits.

The genetic correlation may be reported in a variety of ways. In genetic evaluation of seedstock as performed for a variety of breed associations, this correlation is used to compute Beef Improvement Federation (BIF) accuracy.[2] The BIF accuracy is then reported for each calculated expected progeny difference (EPD). The BIF accuracy, although also ranging from 0 to 1, is always numerically lower than the genetic correlation. This correlation can also be used to calculate reliability or estimate the percent of genetic variation explained. Importantly, all of these metrics – genetic correlations, accuracy, reliability, and percent genetic variance – are all representing the same statistical property of the prediction.

The effect of increased accuracy is a more dependable comparison of the relative genetic merit of selection candidates. As a result, any given selection decision will be more effective and the average genetic merit of selected animals will be greater than if a less-accurate prediction were used. Genomic data simply serves as an additional source of information that can be added to other available data, thus increasing the accuracy of genetic predictions.

Criteria for Selection

There are a large number of genetic predictions that can be used to inform selection decisions. Seedstock genetic evaluations often support predictions for 10 to 15 different traits. These typically include growth traits, carcass traits, calving ease, and reproductive traits. The available information can be used to support a variety of selection priorities. However, the sheer number of traits available can make selection decisions quite difficult to process.

Seedstock producers are generally well versed in selection strategies, and often breed for several different types of animals to meet the needs of their commercial cow-calf customers. For example, it is common to simultaneously breed for cattle with high calving ease and moderate growth potential, and a second line with greater growth and carcass merit. The benefit for the commercial cattlemen is that the seedstock breeder has, in this instance, created 2 general categories of cattle (specifically bulls for use as herd sires), thus simplifying selection decisions.

Another commonly applied approach to simplifying selection is the use of selection indexes. These represent a strategy to provide a single, consensus estimate of genetic merit across a broad range of traits. Indexes are intended to provide comprehensive selection across a range of traits, preventing risky single-trait selection. Selection indexes are developed by defining a production outcome (e.g., calves sold at weaning or premium carcass quality) and then describing the relative contribution of each trait to the economic value of that production outcome. Based on the variation within and among traits, the heritability of the traits, and the economic value of each trait, an optimal combination of relative emphasis can be defined for each trait included in the index. This process removes subjectivity from development of the index and ensures that the ultimate outcome will favorably impact profitability.

Types of Selection Decisions

The emergence of genomic technologies has not altered how genetic improvement is achieved. There are effectively only 3 selection decisions that any producer can make. The first is choosing which animals will be retained / enter the herd. This applies to both replacement females and herd sires. The second selection decision is defining how to assign matings to address individual weaknesses and complement existing strengths. The final selection decision is how many progeny each animal will produce. Elite animals are assigned to produce many progeny, as may be the case with AI sires and donor females.

To execute any of these selection decisions, the breeder/producer needs to be able to rank animals from best to

worst based on their selection criteria. This requires a genetic prediction for economically relevant traits with sufficient accuracy to make the right selection decisions – and avoid mistakes. Genomic technologies, by virtue of their contribution to accuracy, help to ensure that animals are ranked more dependably against their true differences in genetic merit.

The Science Behind Genomics

Genomic technologies provide a mechanism to provide information that complements available pedigree, performance, and progeny data to more accurately inform selection decisions. Given that DNA can be analyzed in every animal very early in life, genetic tests also have the potential to support selection decisions in young animals for a host of traits, including those traits that are difficult to measure or are not expressed until much later in life.

Genetic tests can be designed to support these objectives in 2 ways. The first approach is to develop genetic tests that describe genetic variation in genomic reasons with known associations to phenotypic outcomes. These generally interrogate quantitative trait loci (QTL), regions of the genome linked to quantitative traits like weight, height, or milk production. The challenge with this approach, often referred to as the candidate gene approach, is that the vast majority of economically relevant traits are polygenic, influenced by many regions of the genome. In addition, since the majority of the genetic markers used to query QTL are not causative mutations, but instead markers in close proximity within the genome to unknown causative mutations, the associations may be tenuous and may not apply to animals outside the reference population. Markers that are close to a gene tend to be inherited with that gene, but mutations do occur and identified markers may be segregating differently relative to the causative mutation in different populations of cattle. For this reason, external validation of genomic predictions designed using the candidate gene approach is critically important.

A second and increasingly common approach is to utilize genetic markers that are not selected on the basis of their association with a specific outcome, but based on their ability to describe genetic variation in general. As such, these genetic tests typically include thousands of markers that span the entire genome. This approach seeks to principally evaluate lineage or genomic relatedness and, in so doing, allow inferences to be made regarding an individual animal's genetic merit based on prior knowledge of the genetic potential of other animals possessing similar genomic patterns. This approach is used extensively in seedstock animals to derive genomic information that can be integrated into existing genetic evaluations.

The genome-wide approach typically produces more comprehensive and accurate genomic predictions. However, they require genotypes for a far greater number of genetic markers and therefore are generally more expensive to obtain. A significant recent innovation that is helping to overcome this challenge is imputation. With a sufficient number of higher-density genotypes on the right animals, it is possible to begin to recognize common patterns within the genotypes. Imputation leverages the knowledge of the patterns that are common in the reference population to predict higher-density genotypes from a strategically selected subset of markers. In well documented populations (e.g., Angus, Nelore), the predicted genotypes will correctly match the true genotype in greater than 95% of the markers.[1,3] This degree of imputation accuracy is sufficient to inform predictions with very nearly the same accuracy using a more cost-effective genotyping platform.

Conclusions

Genomic technologies have the potential to provide valuable support to cattle producers' genetic improvement strategies. For veterinarians, there is minimally a responsibility to understand how the technology is applied so as to provide relevant guidance to their clients. The technology is complex, in large part because it is intended to describe the inherently complex. However, there are some key concepts that can be readily understood. Given the value of the technology and the impact it is having in the industry today, it is likely that it will continue to emerge and gain greater adoption throughout the industry. Beef producers, particularly commercial cattlemen, will continue to need sound guidance from their veterinary advisors as they integrate genetic testing and genomic data into their operations.

References

1. Carvalheiro R, Boison SA, Neves HHR, Sargolzaei M, Schenkel FS, Utsunomiya YT, O'Brien AM, Solkner J, McEwan JC, Van Tassell cP, Sonstegard TS, Garcia JF. Accuracy of genotype imputation in Nelore cattle. *Genet Sel Evol* 2014; 46:69.
2. Guidelines for Uniform Beef Improvement Programs. Beef Improvement Federation. 9th ed. 2010.
3. Ventura RV, Lu D, Schenkel FS, Wang Z, Li C, Miller SP. Impact of reference population on accuracy of imputation from 6K to 50K single nucleotide polymorphism chips in purebred and crossbred beef cattle. *J Anim Sci* 2014; 92:1433-1444.